

# Networks and Our Limited Information Horizon

M. Rosvall\*

*Department of Theoretical Physics, Umeå University, 901 87 Umeå, Sweden*

K. Sneppen

*Niels Bohr Institute, Blegdamsvej 17, Dk 2100, Copenhagen, Denmark<sup>†</sup>*

(Dated: February 6, 2008)

In this paper we quantify our limited information horizon, by measuring the information necessary to locate specific nodes in a network. To investigate different ways to overcome this horizon, and the interplay between communication and topology in social networks, we let agents communicate in a model society. Thereby they build a perception of the network that they can use to create strategic links to improve their standing in the network. We observe a narrow distribution of links when the communication is low and a network with a broad distribution of links when the communication is high.

## I. INTRODUCTION

Communication is a fundamental element in maintaining the overall cooperation between different parts of a complex system. Because a complex system consists of many different parts, it matters where signals are transmitted. Thus signaling and traffic is in principle specific, with each message going from an unique sender to a specific target. Networks are therefore a powerful way to represent this constrained communication of the real world [Rosvall & Sneppen, 2003].

We start by using walks in networks with specific targets to quantify the information necessary to locate specific nodes in the network [Sneppen *et al.*, 2005], and also to investigate the constraints limited information sets on the navigability. The process consists of extracting information at the nodes on the walk between a source and a target. The subsequent question is therefore the availability of this information. We therefore let agents in a model society use local communication to self-organize distant communication-pathways. In this way we demonstrate that simple local rules allow agents to build a perception of a dynamic system. This perception guides a targeted signal across the network beyond the information horizon [Friedkin, 1983, Trusina *et al.*, 2005, Valverde & Solé, 2004]. Further we in this minimalistic model find that messages are most effectively forwarded in the presence of hubs with funneling [Jeffrey & Milgram, 1969], like in scale-free networks, while transmission in hub-free networks is more robust against misinformation and failures.

With the locally generated global information of the network, we can take the model society one step further and let the agents use the information to create strategic links. In this way we are able to model the self-organization between communication and topology in social networks, with a feedback between different communication habits and the topology. We observe a narrow distribution of links when the communication is low and a system with a broad distribution of links

when the communication is high.

## II. NAVIGATION IN NETWORKS

The simplest walker is the random walker, which has been used to characterize topological features of networks [Bilke & Peterson, 2001, Monasson, 1999], including first passage times [Noh & Rieger, 2004], large scale modular features [Eriksen *et al.*, 2003], and search using topological features [Adamic *et al.*, 2001]. Here we instead take the opposite approach and consider a direct walker. In particular we quantify the information necessary to locate specific nodes in the network [Sneppen *et al.*, 2005], and investigate the constraints limited information sets on the navigability [Rosvall *et al.*, 2005a].

A walk consists of stepping from node to node via the links between them. The walk from a source node  $s$  to a target node  $t$  may be more or less directed depending on the walkers ability to choose exit links that lead it closer to the target (see Fig. 1(c-d)). We first quantify the information cost in number of bits  $I(s \rightarrow t) = \sum_{j \in \text{path}(st)} \iota_{jt}$  it takes to navigate the shortest path from node  $s$  to node  $t$ , as the sum of the local information  $\iota_{jt}$  on every node  $j$  on a walk “path( $st$ )” leading to target  $t$ . That is,  $\iota_{jt}$  is the number of bits one needs on node  $j$  to select one exit that leads to  $t$ . If no degenerate paths exist, as in Fig. 1(a), then  $\iota_{jt} = \log_2 k_j$ , where  $k_j$  is the degree (number of links) of node  $j$ , since the task is to select one link among  $k_j$ .

When there are two or more degenerate paths from  $j$  to  $t$ , the required information depends on the relative probabilities that one wants to choose each shortest path with, and  $\iota_{jt}$  above generalizes to

$$\iota_{jt} = \log_2(k_j) + \sum_i q_{jit} \log_2 q_{jit}, \quad (1)$$

where  $q_{jit}$  is the probability to choose a link to node  $i$  from node  $j$  on a walk to node  $t$  ( $\sum_i q_{jit} = 1$ ).  $q_{jit} = 0$  if the link is not on the shortest path between  $j$  and  $t$ . We will choose the probability to leave a node along a link on a shortest path between  $s$  and  $t$  to minimize the total information cost  $I(s \rightarrow t)$ . Thus, if there are many degenerate paths, the probability

\*Electronic address: rosvalld@tp.umu.se

<sup>†</sup>URL: <http://cmol.nbi.dk>

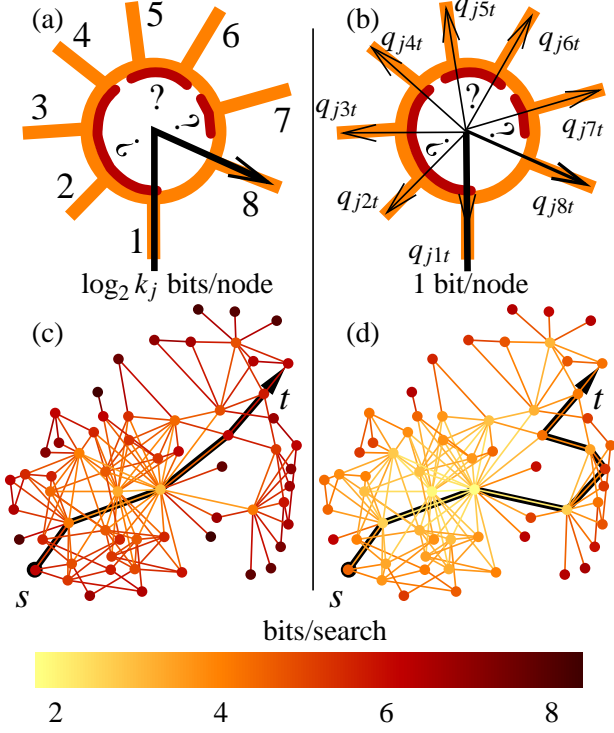


FIG. 1 Navigation in networks. (a) It is possible to ask yes-no questions and successively eliminate groups of wrong exits if the links are ordered. Every yes-no question optimally reduces the number of possible links with  $1/2$  and the cost is  $\log_2 k$  to find the correct exit link. (b) The outcome of the elimination process is uncertain if the information that can be received is limited, and there is a finite probability to make mistakes ( $q_{jit}$  in the direction  $i$  at this node  $j$  on the way to target node  $t$ ). The color of the nodes and links in the lower panel represent the average number of questions necessary to navigate to the node. (c) With in total 13 bits it is possible to find the shortest path between the marked source and target. (d) With the information limited to 1 bit/node the navigation becomes here 3 steps longer (total cost is 8 bits).

to exit to node  $e$  from node  $j$  on the shortest path to  $t$  is

$$q_{jet} = \frac{p_{jet}}{\sum_i p_{jit}}, \text{ where } p_{jit} = \sum_{\text{path}(it)} \prod_{l \in \text{path}(jit)} \frac{1}{k_l} \quad (2)$$

is the probability to walk the shortest path to  $t$  from node  $j$  via the link to node  $i$  in an unbiased walk.

We now turn to the limited information perspective, and assume that the amount of information at a node is limited to  $\iota$  bits (see Fig. 1(b) and Java applet [URL, a]). The walk can now be substantially longer than the actual shortest path. In Fig. 1(d)  $\iota = 1$  and the walk is about 3 steps longer than in Fig. 1(c) with unlimited information. To limit  $\iota_{jt}$  to  $\iota$  we blur the  $q$ -values of node  $j$  in Eq. (2) by a  $\epsilon_{jt} \in [0, \infty]$ , through  $q_{jit} \rightarrow q_{jit}(\epsilon_{jt}) = \frac{q_{jit} + \epsilon_{jt}}{1 + k_j \epsilon_{jt}}$  with  $\epsilon_{jt}$  determined by

$$\iota_{jt} = \log_2(k_j) + \sum_i q_{jit} \log_2 q_{jit} \leq \iota. \quad (3)$$

The navigability of a network is determined by its topology, hence it depends on both the degree distribution and how

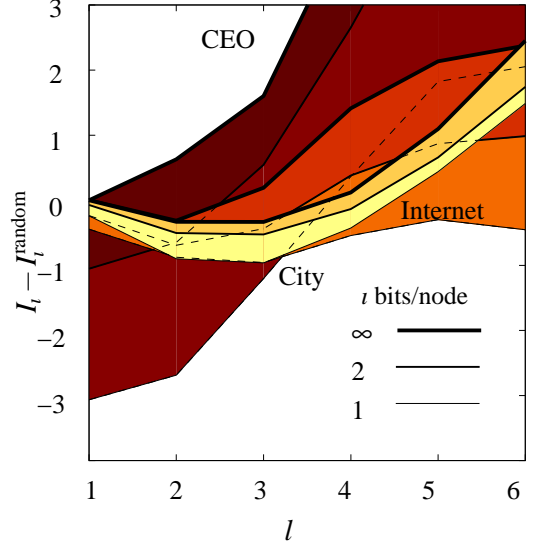


FIG. 2 Information horizon in three real-world networks. We compare information associated to navigation between nodes at distance  $l$ , with the navigation in randomized counterparts (keeping the degree sequence) for  $\iota = \infty, 2$  and  $1$ . The Internet (hardwired Internet of autonomous systems [URL, b]) is more sensitive to limited information than the similarly sized CEO (Chief Executive Officers connected by links if they sit at the same board [Davis & Greve, 1997]). The city network is the Swedish city Malmö with streets mapped to nodes and intersections mapped to links [Rosvall *et al.*, 2005b].

nodes of various degrees are connected to each other. Here we focus on comparing a given real-world network with its randomized counterparts, created by rewiring links such that all nodes conserve their degree, and such that the network remains globally connected [Maslov & Sneppen, 2002]. In Fig. 2 we resolve  $I_t$  into  $I_t(l)$  and examine the average information associated to walking to a specific node a distance  $l$  away in the real and in the randomized ( $I_t^{\text{random}}(l)$ ) network [Trusina *et al.*, 2005].

The pattern that the real networks demand less information than their randomized counterparts on short distance,  $l < 3$ , suggests that many real-world networks favor communication on short distance at the cost of constraining communication beyond this horizon. Furthermore, this feature is more evident with limited than with complete information.

All results until now are based on the assumption that information is available at every node, although the information can be limited. To address the question of how this information can be assembled, we in the next section turn to a social game and let agents in a model society use local communication to build a global perception of the network.

### III. SELF-ASSEMBLY OF INFORMATION IN NETWORKS

To visualize our basic approach we illustrate in Fig. 3 a network composed of individual agents, each of them connected to one or more acquaintances. Each individual communicates with its immediate neighbors to exchange information about

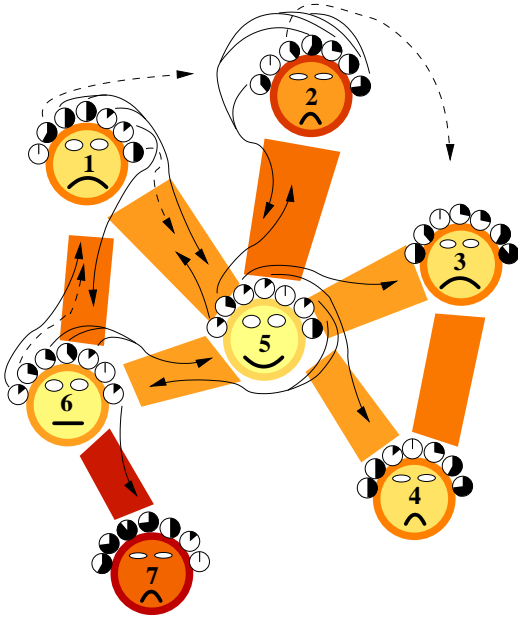


FIG. 3 Modeling self-assembly of information in networks. Agents at nodes, connected by links, communicate with their connected acquaintances about any third target agent in the network, and estimate the quality of the information by its age (clocks over heads correspond to, from left to right, the age of the information about agent 1, 2, ..., 7). The pointers are, for every agent, the acquaintance that connects most efficiently to each of the other agents in the system (here only indicated for agent 1, 2, 5, and 6, dashed pointers are outdated).

agents in other parts of the system. In this way every individual gradually builds up a global perception by knowing people through people [Friedkin, 1982]. In our minimalistic model, we allow each agent to have the information about which of its neighbors that connects most efficiently to each of the other agents in the system. Thus, a perfectly informed agent knows in which direction to send a message to any other agent in the system. If all agents were perfectly informed, any message would be reliably forwarded from sender to recipient, using the information of the subsequent agents along its paths [Jeffrey & Milgram, 1969].

The key question is how different communication rules of the agents influence their possibility to obtain a reliable perception that is robust to dynamical changes of the network. Obviously, the agents need some index of quality that let them judge whom of their acquaintances that has the best knowledge of a particular agent. We have found that the age of the information about an agent gives a very good estimate of the quality [Rosvall & Sneppen, 2006b]. That is, the *perception* consists of, for every agent about any other agent:

- The age of the information about the other agent (clocks in Fig. 3).
- From whom the information came (pointers in Fig. 3).

This defines the model together with the *communication* event:

- Select a random link and let the two agents that it connects communicate about a random third agent. The two agents also update their information about each other.

For example, when an acquaintance of agent 5 in Fig. 3 obtains information about 5, it sets its pointer to 5, and the information starts aging. With successive communication events, the information spreads from agent to agent and gets older and older (we increase the age of all information when all links on average have participated in one communication event). When two agents compare the validity of their pointers to a target agent, like 1 and 6 to 5 in Fig. 3, they validate the newest information as the most correct one. The agent with the oldest information changes its associated pointer to its acquaintance in the communication event, and updates the clock. By letting the agents memorize the acquaintances that provided the newest information about other agents together with the age of this information, they will point in the direction of the fastest communication path from a target. Moreover, the fastest communication paths are typically close to the shortest paths [Rosvall & Sneppen, 2006b].

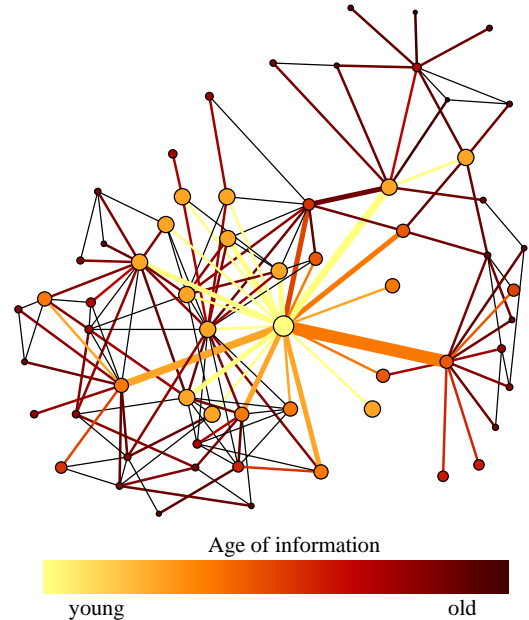


FIG. 4 Self-assembly of information in networks. The size and the color of the nodes reflect the age of the information the well connected agent in the middle has of other agents. The width of the links reflects the relative amount of information they transfer to this agent, and the color the average quality (age) of this information. The agents make use of the hubs to create short communication-paths.

Figure 4 shows the perception around the central node in a model network (see also Java applet [URL, c]). Clearly the information is most up to date in the immediate neighborhood of the agent, but that distant communication-pathways extend the whole network. In a more detailed investigation, we found that messages are most effectively forwarded in the presence of hubs with funneling [Jeffrey & Milgram, 1969], like in scale-free networks, while transmission in hub-

free networks is more robust against misinformation and failures. This is in overall accordance with Stanley Milgram's famous experiment, where letters were transmitted by sequences of acquaintance-acquaintance contacts across USA [Jeffrey & Milgram, 1969, Milgram, 1967]. The choice of acquaintances was based on the participants' network perception, including also geographic closeness of the acquaintance to the target (first steps) and similarity of occupation (later steps) to forward messages [Killworth & Bernard, 1978]. Of course, these are two of many layers that could be added to the model. However, the minimalistic model demonstrates that simple local rules allow agents to build a perception of the system, which is enough to overcome the information horizon [Friedkin, 1983] set by immediate acquaintances. In this way the "small world" is really small [Kochen, 1989, Milgram, 1967], and it makes sense to talk about navigation or search in networks and to quantify the information associated to this process.

Given the agents' perception of the network it is tempting to take this social game one step further, and in the next section we give the agents a social mobility. The agents can thereby get new acquaintances to meet different interests.

#### IV. SELF-ORGANIZATION OF NETWORKS

Social mobility may be seen as the response to the quest for better information access in a social system. We let agents communicate to build a perception of a network as in the previous section, and further allow the agents to use this information to create strategic links. In this way we are able to investigate the feedback between different communication habits and the topology, while the agents self-organize the social network.

The core of the model is the same as in the previous section. To this we add the possibility to rewire the network, and the model can be formulated in the two independent events [Rosvall & Sneppen, 2006a]:

- *Communication*: Select a random link and let the two agents that it connects communicate about a random third agent. The two agents also update their information about each other.
- *Rewiring*: Select a random agent and let it use the local information to ask an acquaintance about whom to establish a link to, to shorten its distance to a randomly chosen other agent (the answer is the agent that the acquaintance points to). Subsequently a random agent loses one of its links.

The communication event is typically repeated of the order of number of links in the system for each rewiring event.

We quantify the self-organization between communication, the network, and the perception of this network, in Fig. 5 (see also Java applet [URL, d]). The perception network is defined by nodes as agents, and links between all pairs of nodes where at least one of the corresponding agents has a pointer to the other agent. Hence the perception network reflects the

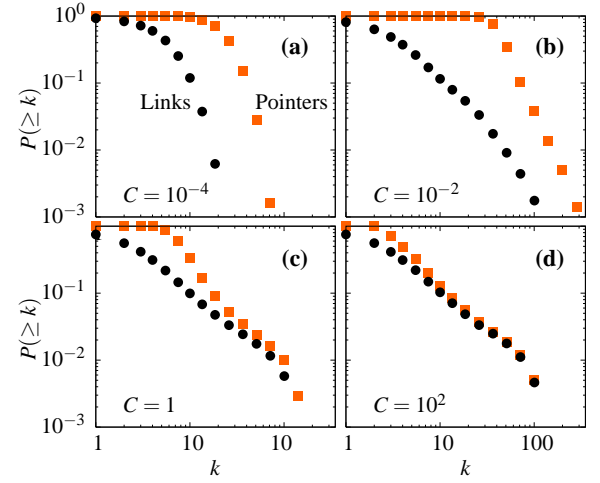


FIG. 5 Illustration of the feedback from communication on the topology of both the communication network and the perception network at four different levels of communication  $C$ .  $C = 1$  corresponds to on average 1 communication event per link and rewiring event. The system size is  $N = 1000$  agents connected by  $L = 2500$  links.

imperfect picture the agents have of their surrounding, as outdated pointers cause diverging communication and perception networks.

We started with random Erdős-Rényi networks [Erdős & Rényi, 1959] (the results are independent of initial conditions) and let the system evolve at different communication levels  $C$ . The system size was  $N = 1000$  agents and  $L = 2500$  links.  $C \cdot L$  is the number of communication events per rewiring event in the network, and the degree  $k$  of a node is its number of links. At low communication level,  $C < 1$ , the perception network has many more links than the communication network, reflecting the failure of agents to perceive connections that are lost recently. As  $C$  approaches  $C \sim 1$  the perception network prunes its links whereas the communication network develops nodes with high degrees (the distribution can be approximated by a power law  $P(k) \propto k^{-2.2}$ ). At even higher values of  $C$  the two networks converge toward the same broad degree-distribution—local communication gives rise to global organization.

The presented model describes a social game where the aim is to be central, and a winner is an agent with many connections that provide short and reliable communication to other agents. The fact that we observe agents with a wide range of degrees reflects the diversity of the possible outcomes of the game, and raises the questions about whether there are some particular strategies with which agents can improve their standing in the network? In a more detailed investigation, we found that individual increase of communication gives both a local gain for the agents that adopt the communication strategy, and a global gain for the whole system [Rosvall & Sneppen, 2006a].

## V. SUMMARY

We started by investigating how the network topology affects the communication ability in various networks, and demonstrated that many real-world networks favors communication on short distance at the cost of constraining communication on long distance. Thereafter we examined the ability to self-organize locally available information in a system such that messages can be guided between distant parts of the network. Our approach was to let agents chat in a model system to self-organize distant communication-pathways. We demonstrated that simple local rules allow agents to build a perception of the system that is robust to dynamical changes and mistakes.

Finally we explored the local dynamic origin of global network organization by modeling the response to information transfer in a simplified social system. We found that a low communication level results in chaotic or Erdős-Rényi like networks, whereas higher communication levels with more reliable communication-pathways lead to structured network topologies with broad degree-distributions.

Overall we have used networks to quantify our limited information horizon. In a society the information horizon is set by each individual's social contacts, which in turn is a part of the global network of human communication. By measuring the information necessary to locate specific nodes in a network we were able to quantify this horizon, and by introducing a social game we investigated different ways to overcome the horizon.

We acknowledge the support of The Danish National Research Foundation: "Models of Life" at NBI.

## References

- [a] Simulation at <http://cmol.nbi/models/bit/bit.html>.
- [b] Website maintained by the NLANR Measurement and Network Analysis Group at <http://moat.nlanr.net>.
- [c] Simulation at <http://cmol.nbi/models/infoflow/infoflow.html>.
- [d] Simulation at <http://cmol.nbi/models/infoflow/infoflow.html>.
- Adamic, L. A., Lukose, R. M., Puniyani, A. R., & Huberman, B. A. [2001] "Search in power-law networks," *Phys. Rev. E* **64**, 46135.
- Bilke, S. & Peterson, C. [2001] "Topological Properties of Citation and Metabolic Networks," *Phys. Rev. E* **46**, 036106.
- Davis, G. F. & Greve, H. R. [1997] "Corporate Elite Networks and Governance Changes in the 1980s," *Am. J. Soc.* **103**, 1.
- Erdős, P. & Rényi, A. [1959] "On random graphs," *Publ. Math. Debrecen* **6**, 290.
- Eriksen, K. A., Simonsen, I., Maslov, S., & Sneppen, K. [2003] "Modularity and Extreme Edges of the Internet," *Phys. Rev. Lett.* **90**, 148701.
- Friedkin, N. E. [1982] "Information Flow Through Strong & Weak Ties in Intraorganizational Social Networks," *Soc. Networks* **3**, 273.
- Friedkin, N. E. [1983] "Horizons of Observability and Limits of Informal Control in Organizations," *Soc. Forces* **62**, 54.
- Jeffrey, T. & Milgram, S. [1969] "An Experimental Study of the Small World Problem," *Sociometry* **32**, 425.
- Killworth, P. & Bernard, H. [1978] "Reverse Small World Experiment," *Soc. Networks* **1**, 159.
- Kochen, M. [1989] *The Small World* (Ablex, Norwood, NJ).
- Maslov, S. & Sneppen, K. [2002] "Specificity and Stability in Topology of Protein Networks," *Science* **296**, 910.
- Milgram, S. [1967] "The Small World problem," *Psychol. Today* **1**, 61.
- Monasson, R. [1999] "Diffusion, Localization and Dispersion Relations on "Small-World" Lattices," *Eur. Phys. J. B* **12**, 555.
- Noh, J. D. & Rieger, H. [2004] "Random Walks on Complex Networks," *Phys. Rev. Lett.* **92**, 118701.
- Rosvall, M., Minnhagen, P., & Sneppen, K. [2005a] "Navigating Networks with Limited Information," *Phys. Rev. E* **71**, 066111.
- Rosvall, M. & Sneppen, K. [2003] "Modeling Dynamics of Information Networks," *Phys. Rev. Lett.* **91**, 178701.
- Rosvall, M. & Sneppen, K. [2006a] "Modeling Self-Organization of Communication and Topology in Social Networks," Submitted.
- Rosvall, M. & Sneppen, K. [2006b] "Self-Assembly of Information in Networks," Submitted.
- Rosvall, M., Trusina, A., Minnhagen, P., & Sneppen, K. [2005b] "Networks and Cities: An Information Perspective," *Phys. Rev. Lett.* **94**, 028701.
- Sneppen, K., Trusina, A., & Rosvall, M. [2005] "Hide and Seek on Complex Networks," *Europhys. Lett.* **69**, 853.
- Trusina, A., Rosvall, M., & Sneppen, K. [2005] "Communication Boundaries in Networks," *Phys. Rev. Lett.* **94**, 238701.
- Valverde, S. & Solé, R. V. [2004] "Internet's Critical Path Horizon," *Eur. Phys. J. B* **38**, 245.